

PATENT ABSTRACTS OF JAPAN

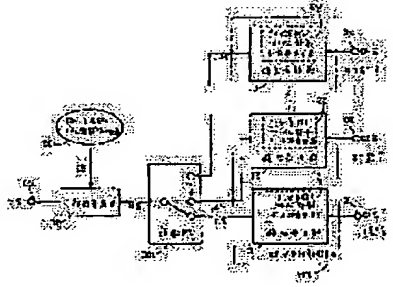
(11)Publication number: 01-211799
(43)Date of publication of application: 24.08.1989

(51)Int.Cl.	G10L 5/00
(21)Application number: 63-037948	(71)Applicant: A T R JIDO HONYAKU DENWA KENKYUSHO:KK
(22)Date of filing: 19.02.1988	(72)Inventor: ABE MASANOBU KUWABARA HISAO KANO KIYOHIRO

(54) REGULAR SYNTHESIZING DEVICE FOR MULTILINGUAL VOICE

(57)Abstract:
PURPOSE: To obtain a voice signal which has individual features as to plural languages by providing a regular synthesizing means which synthesizes the voice signal of a 1st standard speaker with respect to the plural languages and a voice converting means which converts a voice signal outputted by a selecting means into the voice signal of a 2nd speaker to whom individual features are to be added.

CONSTITUTION: A switching part 100 selects one of languages L1WLn, e.g. L1 and outputs the voice signal s11 of a standard speaker A1 of the selected language L1 from a multilingual regular synthesizing group 104 to a voice conversion part 101. The voice quality conversion part 101 receives the voice signal s11, refers to data on the speaker B whose voice is to be given individuality in a voice individual information file 102, and converts the voice signal s11 of the standard speaker A1 into the voice signal s4 of the speaker B, which is outputted. Consequently, the individual features of speaker are given to regularly synthesized voices of respective languages.



PATENT ABSTRACTS OF JAPAN

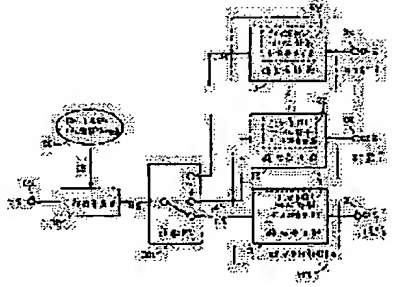
(11)Publication number: 01-211799
(43)Date of publication of application: 24.08.1989

(51)Int.Cl.	G10L 5/00
(21)Application number: 63-037948	(71)Applicant: A T R JIDO HONYAKU DENWA KENKYUSHO:KK
(22)Date of filing: 19.02.1988	(72)Inventor: ABE MASANOBU KUWABARA HISAO KANO KIYOHIRO

(54) REGULAR SYNTHESIZING DEVICE FOR MULTILINGUAL VOICE

(57)Abstract:
PURPOSE: To obtain a voice signal which has individual features as to plural languages by providing a regular synthesizing means which synthesizes the voice signal of a 1st standard speaker with respect to the plural languages and a voice converting means which converts a voice signal outputted by a selecting means into the voice signal of a 2nd speaker to whom individual features are to be added.

CONSTITUTION: A switching part 100 selects one of languages L1WLn, e.g. L1 and outputs the voice signal s11 of a standard speaker A1 of the selected language L1 from a multilingual regular synthesizing group 104 to a voice conversion part 101. The voice quality conversion part 101 receives the voice signal s11, refers to data on the speaker B whose voice is to be given individuality in a voice individual information file 102, and converts the voice signal s11 of the standard speaker A1 into the voice signal s4 of the speaker B, which is outputted. Consequently, the individual features of speaker are given to regularly synthesized voices of respective languages.



⑩ 日本国特許庁(JP) ⑪ 特許出願公開
⑫ 公開特許公報(A) 平1-211799
⑬ Int.Cl.⁴ ⑭ 特許請求の範囲 1 (全6頁)
⑮ G 10 L 5/00 ⑯ 特許請求の範囲 1 (全6頁)
⑰ 特許請求の範囲 1 (全6頁)
⑱ 特許請求の範囲 1 (全6頁)

① 発明の名称 多言語を扱う音声の規則合成装置
② 特 願 昭63-37948
③ 出 願 昭63(1988)2月19日
④ 発 明 者 阿 部 匡 伸 京都府相模原市大字鞍谷小字三平谷5番地 株式会社
エイ・ティ・アール自動車研究所内
⑤ 発 明 者 桑 原 尚 夫 京都府相模原市大字鞍谷小字三平谷5番地 株式会社
エイ・ティ・アール自動車研究所内
⑥ 発 明 者 鹿 野 清 宏 京都府相模原市大字鞍谷小字三平谷5番地 株式会社
エイ・ティ・アール自動車研究所内
⑦ 出 願 人 株式会社エイ・ティ・アール自動車研究所
京都府相模原市大字鞍谷小字三平谷5番地
⑧ 代 理 人 弁護士 深見 久郎 外2名

明 細 書

1. 発明の名称
多言語を扱う音声の規則合成装置
2. 特許請求の範囲
複数の言語のそれぞれについて、規則となる第1の言語の音声単位番号の集合を定め、外面から文字情報番号を受け、前記音声単位番号集合を参照することにより、前記第1の言語の音声番号を合成する複数の規則合成手段と、
前記複数の規則合成手段に接続され、前記複数の言語の中から1つの言語を選択し、選択された言語について前記第1の言語の音声番号を出力する選択手段と、
前記第1の言語の音声、音声に固有の特徴を付与した第2の言語の音声に変換するのに必要な変換番号が登録された変換番号ファイル手段と、
前記選択された言語の合成された前記第1の言語の音声番号を受け、前記変換番号ファイル手段に登録された変換番号に基づいて前記第1の言語の音声番号を前記第2の言語の音声番号に変換する

この発明は、音声の規則合成装置に關し、特に、多言語を扱う規則合成装置に關する。
【従来の技術および発明が解決しようとする課題】
従来、規則合成システムは、各言語ごと(たとえば、日本語、英語、ドイツ語などそれぞれについて)に開発されている。しかし、これらの規則合成システムは、あくまでもネイティブスピーカーが使用することが念頭に置かれ、ネイティブスピーカーでないものが使用することには適していない。すなわち、各言語ごとに開発した規則合成システムは、合成に使用する音声の単位、語彙規則などともにネイティブスピーカーをもとに作成されているため、その規則合成音は、ネイティブスピーカーでないものが話す音声よりも質的な音を出力できるが、一方、これらの規則合成システムの出力可能な音声の質は、1つまたは数語種(男

性の音声、女性の音声、子供らしい音声および老人のような音声などに限定されている。

したがって、規則合成システムがある人々に代わって音声を合成すること、すなわちある人の音声の属性に近い音声を合成し、あたかもその人がその言葉で話しているような音声を出力することができないという課題があった。

その理由として、

(1) 規則合成に用いられる音声単位は、ある話者が発生した数百個の音声から作成されるため、音声の量が大きい。

(2) 音声単位の作成は、全自動で行なうのが難しく、人手がかり、たかきんの話者について音声の単位を作成するのは事実上不可能である。

(3) 規則合成に用いられる音声の単位セットを各話者ごとに作ると、格納しておくメモリ量が膨大なものとなる。

などを挙げることができる。

この発明は、上記のような課題を解決するためになされたもので、各言語について合成された規

- 3 -

る第1の話者の音声信号を規則合成した後で、選択手段により選択された言語について、音声合成手段により合成信号ファイル手段に蓄積された合成信号に就いて、その第1の音声信号を個人的特徴を有する第2の話者の音声信号に置換するので、多言語について個人的特徴を有する音声信号を得ることができる。

【発明の実施例】

第1図は、この発明による多言語を扱う規則合成装置の一実施例を示すブロック図である。

第1図を参照して、この規則合成装置は、入力部101ないし102と、多言語規則合成部104と、切換部100と、声質変換部101と、音声個人情報ファイル102と、出力部103とを含む。

この規則合成装置は、言語11ないし言語1NのN個の言語（たとえば、英語、ドイツ語、中国語、日本語など）を扱うことができる。

多言語規則合成部104は、各言語について文字情報信号が与えられる入力部からの信号をそれぞれ処理する規則合成部を含む。たとえば、ある

- 5 -

規則合成器に別声者の個人的特徴を持たせることを目的とする。

【課題を解決するための手段】

この発明に係る多言語を扱う音声の規則合成装置は、複数の言語のそれぞれについて外部から文字情報信号を受け、各言語について標準となる第1の話者の音声単位信号集合を参照することにより、各言語の第1の話者の音声信号を合成する装置、各言語の第1の話者の第1の標準の音声信号のうちいずれかを選択して出力する選択手段と、第1の話者の音声を、音声に個人的特徴を付与したい第2の話者の音声に置換するのに必要な変換信号を蓄積した変換信号ファイル手段と、選択手段から出力された第1の話者の音声信号と、変換信号ファイル手段に蓄積された変換信号に基づいて、第2の話者の音声信号に置換する音声変換手段とを含む。

【作用】

この発明における多言語を扱う音声の規則合成装置は、規則合成手段によりその言語の標準とな

- 4 -

る言語11について、入力部101に、外部から言語11に関する文字およびアクセントなどの文字列ならびに標準信号を含む文字情報信号101が与えられる。言語11の標準の話者A1の音声の単位セット情報ファイル102を含む規則合成部101は、文字情報信号101を受け、音声の単位セット情報ファイル102を参照して、その言語の標準の話者A1の音声信号101を規則合成する。一方、言語11以外の他の言語についても、各規則合成部21ないし1N1によって、同様に、各言語の標準の話者A2ないしAnの音声信号21ないしn1を規則合成することができ。

ここで、音声の単位セット情報ファイル102ないし1N2は、各言語11ないし1Nの標準の話者A1ないしAnについて、それぞれ、言葉や音節などの音声の単位となる情報が予め蓄えられたデータベースである。

切換部100は、各言語11ないし1Nのうちから、いずれか1つの言語、たとえば、言語11を選択し、多言語規則合成部104から出力され

- 6 -

英語の規則合成部11で合成されたあるアメリカ人話者A1の音声信号11を、個人性を付与したい日本人Bの音声信号4に置換する場合について説明する。

第2図は、変換コードブックの作成手順を示すフロー図である。

第2図を参照して、以下に変換コードブック41、42、43を求める手順について説明する。まず、ステップ301および302において、アメリカ人話者A1および日本人話者Bに同一の単語を音声させ、それぞれの音声にLPC分析を施し、パワー、ビッチ周波数およびスペクトルパラメータを求める。次に、ステップ303および304において、スペクトルパラメータをベクトル量子化し、ステップ305および306でパワーをスカラ量子化し、ステップ307および308においてビッチ周波数をスカラ量子化する。話者Aおよび話者Bの音声の時間に対応をとるために、スペクトルパラメータを用いて、ステップ309においてDouble Split

- 8 -

403において話者A1のパワーコードブック403のパワーがスカラ量子化され、ステップ404において話者A1のビッチ周波数コードブック404のビッチ周波数がスカラ量子化される。これらのビッチ周波数がスカラ量子化される。これらの量子化されたパラメータを復号化する過程において、前述の変換コードブック41、42、43が使用される。すなわち、ステップ405において、話者A1から話者Bへのスペクトル変換コードブック41を用い、ステップ406において、パワー変換コードブック43を用い、ステップ407においてビッチ周波数変換コードブック42を用いる。そして、変換された各パラメータを用いてステップ408で話者Bの音声合成される。これによって、日本人話者Bの声をもった英語の規則合成装置が得られる。

第4図は、この発明による多言語を扱う規則合成装置を含む規則合成システムのハードウェア構成を示すブロック図である。

第4図を参照して、この規則合成システムは、アンプ1とローパスフィルタ2とA/D変換器3

- 10 -

とコンピュータシステム4を含む。アンプリは入力された音声信号を増幅するものであり、ローパスフィルタ2は増幅された音声信号から折返し雑音除去するものである。A/D変換器3は音声信号をデジタル化するものである。A/D変換器は16ビットのデジタル信号に変換するものである。コンピュータ4は、構成要素7と（構成要素8）5と磁気ディスク6と中央処理7とプリンタ8とを含む。この時間による多言語環境学習用の構成要素は、第1図の順列合成装置5内において構成される。

以上に述べた多くの書語を扱う規則合成装置は、特に、多数の書語をある話者が発する代わりに音素を出力する多音節語彙代用用規則合成システムとして有用である。

【既知の所見】

以上のように、この説明によれば、多書院につ

一一一

し A_n の音声番号、s 3 は話者 B の音声個人情報
番号、s 4 は話者 B の番号を示す。

東京証券取引所 上場銘柄

自動鋼板試験研究所

代理人 舟越士 久郎
(ほか2名)

- 13 -

-1180-

-1181-

号に要換する音声変換手段とを含むので、多音語
について團人的特徴を有する音声信号を得ること
ができる。

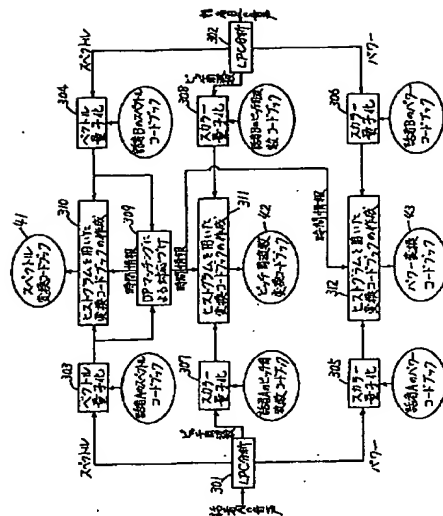
4. 図面の簡単な説明

第1図は、この発明による多要素を扱う規則台成装置の一次規則を示すブロック図である。第2図は、在庫コードブックの作成手順を示すフロー図である。第3図は、戸建住宅の建物種別を算定する手順を示すフロー図である。第4図は、この発明による多要素を扱う規則台成装置を含む規則台成システムのパードウェア構成を示すブロック図である。

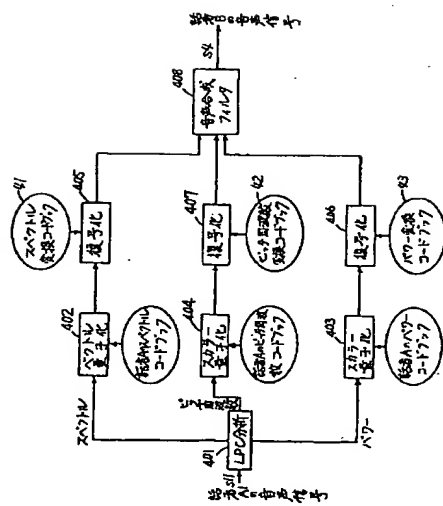
図において、1はアンプ、2はローパスフィルタ、3はA/D変換器、4はコンピュータシステム、5は制御台成装置、10ないし100は入力部、11ないし111は検測台成部、12ないし121は警報発生単位情報フィードバック部、101は警報発生部、102は音声入力情報フィードバック部、103は出力部、104ないし109は文字情報信号、11ないし111は読者A1ない

- 12 -

第2圖



第 3 圖



第 4 圖

